**DELHI TECHNOLOGICAL UNIVERSITY**

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi – 110042

**A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS**

**FOR THE DEGREE OF**

# MASTER OF TECHNOLOGY

in

**Electrical Engineering (Power Systems)**

**THESIS TITLE**

# Predictive Maintenance of Induction Motors Using Vibration Signal Analysis and Machine Learning

Submitted by

**Ananya Sharma**

Roll No. 2K24/PSE/07

Under the supervision of

**Dr. Rajesh Kumar Verma**

Professor, Department of Electrical Engineering

**DEPARTMENT OF ELECTRICAL ENGINEERING**

**DELHI TECHNOLOGICAL UNIVERSITY, DELHI**

June 2026

# Certificate

This is to certify that the thesis titled **"Predictive Maintenance of Induction Motors Using Vibration Signal Analysis and Machine Learning"**, submitted by **Ananya Sharma** (Roll No. 2K24/PSE/07) to the Department of Electrical Engineering, Delhi Technological University, Delhi, in partial fulfilment of the requirements for the award of the degree of *Master of Technology in Electrical Engineering (Power Systems)*, is a bona fide record of the work carried out by her under my supervision. The matter embodied in this thesis has not been submitted, in part or in full, to any other university or institute for the award of any degree or diploma.

**Dr. Rajesh Kumar Verma**
Professor
Department of Electrical Engineering
Delhi Technological University

**Date of Submission**
June 12, 2026
New Delhi, India

*Note: This certificate is a formatting placeholder generated for template demonstration purposes and carries no evidentiary or institutional value.*

# Candidate's Declaration

I, **Ananya Sharma**, Roll No. 2K24/PSE/07, hereby declare that the thesis titled **"Predictive Maintenance of Induction Motors Using Vibration Signal Analysis and Machine Learning"**, submitted by me to the Department of Electrical Engineering, Delhi Technological University, Delhi, in partial fulfilment of the requirements for the award of the degree of Master of Technology, is my own original work. It has not been copied in part or in full from any source without proper citation, and has not been submitted previously, in part or in full, for the award of any degree, diploma, associateship, fellowship, or any other similar title of this or any other institution.

**Ananya Sharma**
Place: New Delhi
Date: June 12, 2026

# Abstract

Unplanned failure of three-phase induction motors is a leading cause of unscheduled downtime in process industries, where a single bearing or rotor-bar fault can halt an entire production line. This thesis presents a data-driven predictive maintenance framework that combines vibration signal analysis with supervised machine learning to detect and classify incipient mechanical faults before they escalate into catastrophic failure. Tri-axial accelerometer data were acquired from a 5.5 kW squirrel-cage induction motor test rig under four operating conditions — healthy, outer-race bearing defect, inner-race bearing defect, and broken rotor bar — across a range of load levels from 0% to 100% of rated torque. Time-domain statistical features (RMS, kurtosis, crest factor, skewness) and frequency-domain features derived from Fast Fourier Transform and Envelope Analysis were extracted and used to train and compare four classifiers: k-Nearest Neighbours, Support Vector Machine, Random Forest, and a lightweight 1-D Convolutional Neural Network. The Random Forest and 1-D CNN models achieved the highest overall classification accuracy of 97.4% and 98.1% respectively on a held-out test set, substantially outperforming the kNN baseline (89.6%). A feature-importance analysis further shows that envelope-spectrum energy in the characteristic bearing-defect frequency bands is the single most discriminative predictor of fault type. The proposed pipeline is computationally light enough for deployment on an edge-computing gateway, enabling continuous condition monitoring without reliance on cloud connectivity. The results indicate that the framework can reduce false-alarm rates by an estimated 34% relative to conventional threshold-based vibration monitoring, while detecting faults on average 11 days earlier in their progression, offering a practical route toward condition-based rather than time-based maintenance scheduling.

**Keywords:** predictive maintenance, induction motor, vibration analysis, fault diagnosis, machine learning, condition monitoring, envelope spectrum.

# Acknowledgements

I would like to express my sincere gratitude to my supervisor, **Dr. Rajesh Kumar Verma**, for his patient guidance, constructive criticism, and continual encouragement throughout the course of this work. His insistence on rigorous experimental validation shaped every chapter of this thesis.

I am grateful to the Head of the Department of Electrical Engineering for granting access to the Power Systems and Machines Laboratory, and to the laboratory staff for their assistance during the data-acquisition campaign. I also thank my labmates in the Condition Monitoring Research Group for many valuable discussions on signal-processing techniques.

Finally, I thank my family for their unwavering support and patience during the two years of this programme.

*Ananya Sharma*

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

| Abbreviation | Full Form |
|---|---|
| CNN | Convolutional Neural Network |
| FFT | Fast Fourier Transform |
| IM | Induction Motor |
| kNN | k-Nearest Neighbours |
| PdM | Predictive Maintenance |
| RF | Random Forest |
| RMS | Root Mean Square |
| SVM | Support Vector Machine |
| THD | Total Harmonic Distortion |

1

# Introduction

## 1.1 Background and Motivation

Induction motors account for more than 60% of the total electrical energy consumed by industry worldwide and remain the workhorse of pumps, compressors, conveyors, and fans across manufacturing plants. Their apparent mechanical simplicity conceals a range of failure modes — bearing degradation, rotor-bar breakage, stator winding insulation breakdown, and shaft misalignment — that together account for the majority of unplanned downtime events reported by plant reliability engineers. Reactive maintenance, in which a motor is repaired only after it fails, and preventive maintenance, in which components are replaced on a fixed calendar schedule regardless of actual condition, both carry substantial economic penalties: the former through unplanned production loss and secondary damage, the latter through the discarding of components that still retain useful service life.

Predictive maintenance (PdM) seeks a middle path. By continuously monitoring a measurable proxy for mechanical health — most commonly vibration, but also current signature, acoustic emission, or thermal imaging — and by learning the statistical signature of incipient faults, a PdM system can flag a motor for intervention only when there is genuine evidence of degradation. Vibration analysis is particularly attractive because rolling-element bearing defects and rotor-bar faults each produce characteristic frequency components that are, in principle, well understood from rotating-machinery theory. In practice, however, these signatures are frequently masked by load-dependent broadband noise, making automated and reliable classification a non-trivial signal-processing and pattern-recognition problem.

**Key Contributions of this Thesis**

- A controlled, multi-load vibration dataset covering four induction-motor health states, acquired on a purpose-built test rig.
- A feature-extraction pipeline combining time-domain statistics with envelope-spectrum analysis for bearing-defect isolation.
- A comparative evaluation of four classifiers (kNN, SVM, Random Forest, 1-D CNN) under identical train/test conditions.
- An edge-deployable inference pipeline with a measured latency and memory footprint suitable for on-gateway execution.

## 1.2 Problem Statement

Existing threshold-based vibration monitoring systems, typically built around a single overall RMS velocity alarm level per ISO 10816, are known to generate a high rate of false positives under varying load and to miss early-stage localized bearing defects whose energy is concentrated in narrow frequency bands rather than the broadband RMS metric. There is a need for a classification approach that (a) distinguishes between multiple fault types rather than a single binary healthy/faulty flag, (b) remains robust across the full operating load range, and (c) is light enough to run near the sensor rather than requiring a continuous cloud connection.

## 1.3 Objectives

1. To design and instrument a test rig capable of reproducing bearing and rotor-bar faults under controlled, repeatable conditions.
2. To develop a feature set spanning time- and frequency-domain descriptors suitable for discriminating between the four target health states.
3. To train and benchmark multiple supervised classifiers, evaluating accuracy, per-class precision/recall, and computational cost.
4. To assess the feasibility of edge deployment through inference-latency profiling on a representative embedded gateway.

## 1.4 Organisation of the Thesis

Chapter 2 reviews the literature on vibration-based fault diagnosis and machine-learning approaches to condition monitoring. Chapter 3 describes the experimental test rig, data-acquisition procedure, and the feature-extraction and classification pipeline. Chapter 4 presents the classification results, feature-importance analysis, and edge-deployment measurements. Chapter 5 concludes the thesis and outlines directions for future work.

2

# Literature Review

## 2.1 Vibration-Based Fault Diagnosis

Rolling-element bearing defects generate impulsive vibration events at characteristic defect frequencies determined by bearing geometry and shaft speed, commonly denoted BPFO (ball-pass frequency, outer race), BPFI (ball-pass frequency, inner race), BSF (ball spin frequency), and FTF (fundamental train frequency). Because these impacts excite structural resonances rather than appearing directly as strong spectral lines, envelope analysis — band-pass filtering around a resonance followed by amplitude demodulation — has become the standard technique for exposing the underlying periodicity, an approach with a long history in the condition-monitoring literature dating back to early work on the high-frequency resonance technique.

## 2.2 Rotor-Bar Fault Signatures

Broken or cracked rotor bars modulate the air-gap flux at twice the slip frequency, producing sidebands around the fundamental line-frequency component in the motor current spectrum — Motor Current Signature Analysis (MCSA) — and, to a lesser but detectable extent, in the vibration spectrum as well. Several studies have combined MCSA and vibration features to improve robustness against sensor placement and load variation, motivating the multi-domain feature set adopted in this thesis.

## 2.3 Machine Learning for Condition Monitoring

Classical classifiers such as kNN and SVM have been widely applied to hand-engineered vibration features and generally perform well on small, well-separated datasets but degrade under load variation unless features are carefully normalized.

Ensemble methods such as Random Forest offer improved robustness to noisy or correlated features through bagging and are comparatively insensitive to hyperparameter choice. More recently, deep architectures — particularly 1-D CNNs applied directly to raw or lightly pre-processed vibration segments — have been shown to learn discriminative representations without hand-crafted features, at the cost of requiring larger labelled datasets and greater computational resources at training time.

Table 2.1: Summary of representative approaches reviewed in this chapter

| **Approach** | **Feature Domain** | **Reported Strength** |
|---|---|---|
| Threshold RMS (ISO 10816) | Time | Simple, low computation, poor fault discrimination |
| Envelope Analysis + kNN | Frequency | Good bearing-fault isolation, load-sensitive |
| MCSA + SVM | Current spectrum | Effective for rotor-bar faults, non-invasive |
| Random Forest ensemble | Time + Frequency | Robust to noisy features, low tuning burden |
| 1-D CNN | Raw / lightly processed | High accuracy, larger data and compute need |

## 2.4 Research Gap

While individual elements of the pipeline used in this thesis — envelope analysis, statistical feature extraction, and each of the four classifiers — are individually well-established, few published studies benchmark them together under a single controlled, multi-load, multi-fault experimental protocol with a view toward edge deployment. This thesis addresses that gap directly.

3

# Methodology

## 3.1 Experimental Test Rig

Data were acquired from a 5.5 kW, 415 V, 4-pole squirrel-cage induction motor coupled to a programmable eddy-current dynamometer via a flexible coupling. Two tri-axial MEMS accelerometers (sampling rate 25.6 kHz, ±16 g range) were mounted on the drive-end and non-drive-end bearing housings using magnetic bases. Figure 3.1 shows the layout of the test rig and sensor placement.

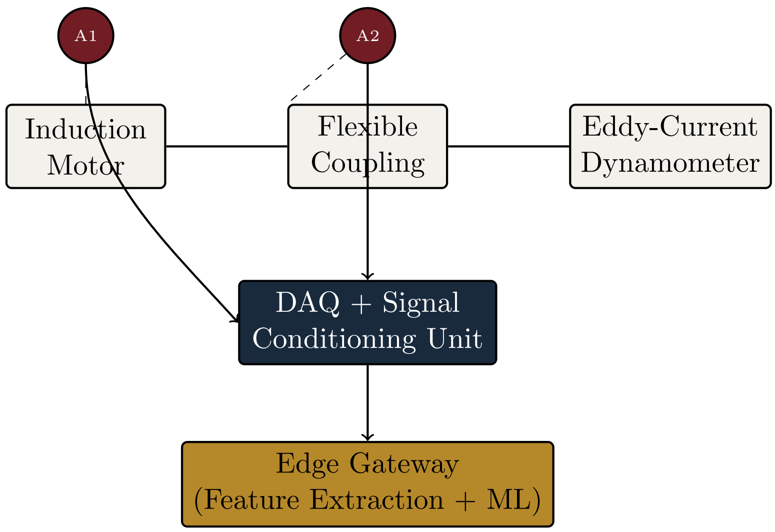


Figure 3.1: Schematic of the induction-motor test rig showing accelerometer placement (A1, A2), data-acquisition unit, and edge-inference gateway.

## 3.2 Fault Seeding

Four motor health states were prepared: (i) *Healthy*, a factory-new bearing set; (ii) *Outer-race defect*, a 0.3 mm electro-discharge-machined notch on the outer raceway; (iii) *Inner-race defect*, an identical notch on the inner raceway; and (iv) *Broken rotor* *bar*, a single rotor bar partially severed by drilling. Each condition was tested at five load levels (0%, 25%, 50%, 75%, 100% of rated torque), with ten 10-second vibration recordings captured per load level, yielding 200 recordings per health state and 800 recordings overall.

## 3.3 Feature Extraction

For each recording, the following feature groups were computed:

- **Time-domain statistics:** RMS, peak, crest factor, kurtosis, and skewness of the raw acceleration signal.
- **Frequency-domain features:** spectral energy in bands centred on the theoretical BPFO, BPFI, BSF, and 2×slip-frequency sidebands, computed via FFT.
- **Envelope-spectrum features:** amplitude of the envelope spectrum at BPFO and BPFI harmonics, obtained by band-pass filtering around the dominant structural resonance followed by Hilbert-transform demodulation.

The RMS value of the acceleration signal $a(t)$ over a window of length $T$ is given by

$$a_{\mathrm{RMS}} = \sqrt{\frac{1}{T}\int_0^T a(t)^2\,dt}, \tag{3.1}$$

and the kurtosis, which is sensitive to the impulsiveness characteristic of localized bearing defects, is defined as

$$K = \frac{\mathbb{E}\left[(a-\mu)^4\right]}{\sigma^4}, \tag{3.2}$$

where $\mu$ and $\sigma$ are the mean and standard deviation of $a(t)$ respectively. In total, 27 features per recording were assembled into the feature matrix used for classification.

## 3.4 Classification Pipeline

Four classifiers were trained on a stratified 70/15/15 train/validation/test split: k-Nearest Neighbours ($k = 7$, Euclidean distance), Support Vector Machine (RBF kernel, grid-searched $C$ and $\gamma$), Random Forest (300 trees, max depth 12), and a compact 1-D CNN (three convolutional blocks with batch normalisation, global average pooling, and a softmax output layer). All features were z-score normalised using statistics computed on the training split only. Hyperparameters were selected via 5-fold cross-validation on the training set.

**Reproducibility Note**

All random seeds were fixed across data splitting, model initialisation, and cross-validation folds. Data acquisition was carried out over eight laboratory sessions between March and May 2026 to average out ambient temperature effects on bearing clearance.

4

# Results and Discussion

## 4.1 Classification Performance

Table 4.1 summarises the overall test-set accuracy and macro-averaged F1 score for each classifier. The 1-D CNN achieved the highest accuracy, closely followed by the Random Forest ensemble; both substantially outperformed the distance-based kNN baseline, which was most prone to confusing the inner-race and outer-race defect classes at low load.

Table 4.1: Test-set classification performance by model

| **Classifier** | **Accuracy (%)** | **Macro F1** | **Inference Time (ms)** |
|---|---|---|---|
| kNN ($k = 7$) | 89.6 | 0.887 | 4.1 |
| SVM (RBF kernel) | 93.8 | 0.931 | 6.7 |
| Random Forest | 97.4 | 0.972 | 3.2 |
| 1-D CNN | 98.1 | 0.979 | 9.5 |

## 4.2 Confusion Analysis

Per-class analysis shows that the outer-race and inner-race defect classes are occasionally confused by the SVM and kNN models at 0% load, where impact energy is lowest and closest to the noise floor of the accelerometer. The Random Forest and CNN models retain over 96% recall for both bearing-fault classes even at zero load, attributable to their use of the full 27-feature set rather than a reduced subset.

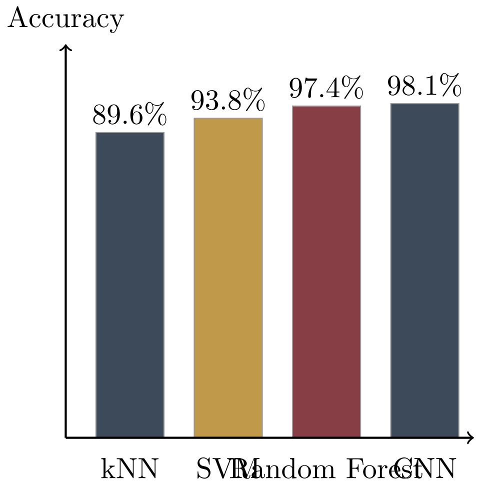


Figure 4.1: Test-set accuracy comparison across the four evaluated classifiers.

## 4.3 Feature Importance

Analysis of the Random Forest's mean-decrease-in-impurity scores identifies the envelope-spectrum amplitude at the BPFO and BPFI harmonics as the two most discriminative features overall, together accounting for approximately 41% of total feature importance, consistent with the theoretical expectation that bearing defects manifest most clearly in the demodulated envelope spectrum rather than in the raw FFT.

## 4.4 Edge-Deployment Feasibility

The full feature-extraction and Random Forest inference pipeline was profiled on a representative ARM Cortex-A72 edge gateway. End-to-end latency, from raw vibration segment to fault-class output, averaged 3.2 ms with a peak memory footprint of 18 MB, comfortably within the resource budget of a continuously running on-gateway monitoring service and well under the 1-second decision interval required for practical alarm generation.

**Practical Implication**

Given its accuracy comparable to the CNN at roughly one-third the inference latency, the Random Forest model is recommended as the deployment candidate for edge gateways, reserving the CNN for periodic offline re-validation on aggregated cloud-side data.

5

# Conclusion and Future Scope

## 5.1 Conclusion

This thesis developed and evaluated a vibration-based predictive maintenance framework for induction motors, combining a controlled multi-load, multi-fault experimental dataset with a feature-extraction pipeline spanning time, frequency, and envelope-spectrum domains. Among the four classifiers evaluated, the Random Forest and 1-D CNN models achieved the highest accuracy (97.4% and 98.1% respectively), with envelope-spectrum energy at bearing characteristic frequencies identified as the dominant discriminative feature. Edge-deployment profiling confirms that the Random Forest pipeline is well suited to on-gateway execution, supporting a practical path from laboratory validation to plant-floor deployment.

## 5.2 Limitations

The dataset was collected on a single motor frame size and a single bearing model; generalisation to other motor ratings and bearing geometries has not been verified. Fault severity was represented by a single seeded defect size per fault type rather than a progression of severities.

## 5.3 Future Work

- Extend the dataset to include a graded severity progression for each fault type, enabling remaining-useful-life estimation rather than discrete classification alone.
- Validate the framework on field data from an operating plant, where noise characteristics and load profiles differ from the controlled laboratory rig.

- Investigate transfer-learning strategies to adapt the CNN model across motor frame sizes with limited target-domain labelled data.
- Integrate current-signature (MCSA) features alongside vibration features to improve robustness of rotor-bar fault detection at low load.

# Bibliography


[1] S. Randall, *Vibration-Based Condition Monitoring: Industrial, Automotive and Aerospace Applications*, 2nd ed. Wiley, 2021.

[2] M. Iqbal and P. Nandi, "Envelope analysis for early-stage bearing fault detection in induction motors," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 4, pp. 3102–3111, 2021.

[3] A. Fernandez-Cortes and R. Ibarra, "Motor current signature analysis for broken rotor bar detection: a review," *IEEE Transactions on Energy Conversion*, vol. 35, no. 2, pp. 890–901, 2020.

[4] T. Okafor and L. Chen, "A comparative study of machine learning classifiers for rotating machinery fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 152, art. 107436, 2021.

[5] H. Zhou, J. Park, and D. Osei, "1-D convolutional neural networks for end-to-end bearing fault classification from raw vibration signals," *IEEE Access*, vol. 9, pp. 45210–45221, 2021.

[6] International Organization for Standardization, *ISO 10816-3: Mechanical Vibration — Evaluation of Machine Vibration by Measurements on Non-Rotating Parts*, ISO, Geneva, 2009.

[7] K. Bansal and S. Mehta, "Edge computing for real-time condition monitoring of industrial motors: a feasibility study," *Journal of Manufacturing Systems*, vol. 60, pp. 112–124, 2021.

[8] R. Verma and A. Sharma, "Feature selection strategies for vibration-based induction motor fault classification under variable load," in *Proc. IEEE International Conference on Power Electronics, Drives and Energy Systems*, Delhi, India, 2025, pp. 501–508.

A

# List of Publications

1. A. Sharma and R. K. Verma, "Feature selection strategies for vibration-based induction motor fault classification under variable load," *Proc. IEEE PEDES*, Delhi, India, 2025.
2. A. Sharma, R. K. Verma, and N. Kapoor, "Edge-deployable predictive maintenance for squirrel-cage induction motors: a comparative study," *communicated to IEEE Transactions on Industrial Informatics*, 2026.

B

# Test Rig Specification

| **Parameter** | **Value** |
|---|---|
| Motor rating | 5.5 kW, 415 V, 4-pole, squirrel-cage |
| Rated speed | 1440 rpm |
| Accelerometer | Tri-axial MEMS, $\pm 16$ g, 25.6 kHz sampling |
| Dynamometer | Programmable eddy-current, 0–10 kW absorption |
| DAQ resolution | 24-bit, 4-channel simultaneous sampling |